DOCUMENT RESUME

ED 404 349 TM 026 107

AUTHOR	Thompson, Bruce; Melancon, Janet G.
TITLE	Using Item "Testlets"/"Parcels" in Confirmatory Factor Analysis: An Example Using the PPSDQ-78.
PUB DATE	Nov 96
NOTE	27p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Tuscaloosa, AL, November, 1996).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Statistical Distributions; *Structural Equation Models; *Test Construction
IDENTIFIERS	*Confirmatory Factor Analysis; Item Parcels; Personal Preferences Self Description Quest; *Testlets

ABSTRACT

This study investigated the benefits of creating item "testlets" or "parcels" in the context of structural equation modeling confirmatory factor analysis (CFA). Testlets are defined as groups of items related to a single content area that is developed as a unit. The strategy is illustrated using data from the administration of the Personal Preferences Self-Description Questionnaire (PPSDQ-78). Testlets or item "parcels" were empirically created by combining items to create score aggregates that could be subjected to CFA analysis. Augmenting analyses by creating testlets can provide more complete understanding regarding the quality of a structural model by altering the measurement model while leaving the structural model intact. The heuristic example of testlet creation involves responses by 422 undergraduates to the PPSDQ. Items were combined through five iterations to yield data that were progressively more normally distributed. Five appendixes present descriptive statistics for the five iterations. The creation of testlets or parcels is recommended whenever data depart substantially from the distributional assumptions of CFA. (Contains 5 tables and 26 references.) (Author/SLD)

ppsdq6ms.wp1 9/28/96

USING ITEM 'TESTLETS'/'PARCELS' IN CONFIRMATORY FACTOR ANALYSIS: AN EXAMPLE USING THE PPSDQ-78

Bruce Thompson

Texas A&M University
and
Baylor College of Medicine

Janet G. Melancon

Loyola University (LA)

Paper presented at the annual meeting of the Mid-South Educational Research Association, Tuscaloosa, AL, November 7, 1996.

ABSTRACT

The present study investigated the benefits of creating item "testlets" or "parcels" in the context of structural equation modeling confirmatory factor analysis. The strategy is illustrated using data from an administration of the <u>Personal Preferences Self-Description Questionnaire</u> (PPSDQ-78). Augmenting analyses by creating "testlets" can provide more complete understanding regarding the quality of a structural model, by altering the measurement model while leaving the structural model intact.

Views regarding the nature of validity have evolved continuously during the last 50 years. This evolution has been chronicled by various scholars (cf. Cronbach, 1989; Moss, 1992; Shepard, 1993). Recent thinking is reflected in contemporary professional standards.

These various sets of professional standards recognize that validity (and reliability) are not properties of tests, and consequently that it is fully inappropriate to describe a test as being valid (or reliable), or to speak of the validity (or reliability) of a test. Informed journal editorial guidelines incorporate these views (Thompson, 1994).

Regarding validity, the AERA/APA/NCME test standards indicate that validity evaluation requires evidence supporting "the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores" (AERA/APA/NCME, 1985, p. 9). Similarly, the personnel evaluation standards of the Joint Committee on Standards for Educational Evaluation (1988) stated that:

> Valid means that what was intended to be measured was measured. Specifically here, valid refers to the degree to which evidence supports the inferences that are drawn from the measurement instruments or procedures. Valid does not refer to the instruments or procedures themselves. Thus, a particular measure may be valid for one purpose but have little or no validity for another purpose. (p. 98)

Finally, the program evaluation standards of the Joint Committee (1994) have been certified by ANSI as the American National Standards. Regarding validity, these standards state that validity "concerns the soundness or trustworthiness of the inferences that are made from the results of the information gathering process" (p. 145). Validation is "the process of compiling evidence that supports the interpretations and uses of the data and information collected using one or more of these instruments and procedures" (p. 145).

Recent views of validity emphasize the role of falsification (Popper, 1962), and that a measure may not be deemed credible until the measure has survived serious efforts to disconfirm the integrity of the measure's scores. As Moss (1995) explained,

> A "strong" program of construct validation requires an explicit conceptual framework, testable hypotheses deduced from it, and multiple lines of relevant evidence to test the hypotheses. Construct validation is most efficiently guided by the test of "plausible rival hypotheses" which suggests credible alternative explanations or meanings for the test score that are challenged and refuted by the evidence collected... Essentially, test validation examines the fit between the meaning of the test score and the measurement intent, whereas construct validation entails the evaluation of an entire theoretical framework. (pp. 6-7)

Factor analysis has long been associated with efforts to evaluate construct validity. For example, historically "construct validity has [even] been spoken of as... 'factorial validity'" (Nunnally, 1978, p. 111). Joy Guilford's article some 50 years ago is illustrative:

> The *factorial validity* of a test is given by its loadings in meaningful, common, reference factors. This is the kind of validity that is really meant when the question is asked "Does this test measure what it is supposed to measure?" A more pertinent question should be "What does this test measure?" The answer then should be in terms of factors and their loadings [sic]... I predict a time when any test author will be expected to present information regarding the factor composition of his [sic] tests. (Guilford, 1946, pp. 428, 437-438, emphasis added)

Similarly, Gorsuch (1983, p. 350) has noted that, "A prime use of factor analysis has been in the development of both the operational constructs for an area and the operational representatives for the theoretical constructs." In short, "factor analysis is intimately involved with questions of validity.... Factor analysis is at the heart of the measurement of psychological constructs" (Nunnally, 1978, pp. 112-113).

However, three basic factor analytic methods can be employed in evaluating score integrity: (a) exploratory factor analysis (CFA) (cf. Gorsuch, 1983); (b) confirmatory rotation of factors

extracted using exploratory methods (Thompson, 1992); and confirmatory factor analysis (CFA) (cf. Jöreskog & Sörbom, 1989). Regarding exploratory EFA, Thompson and Daniel (1996) noted:

> EFA (as against CFA) is not readily amenable to testing rival hypotheses and falsifying models, because in EFA an expected structure either emerges or doesn't. Essentially, in EFA a single model is tested, and rival models are not. Of course, such evidence is useful, particularly when cumulated across studies, but is inherently limited. (p. 204)

CFA requires the researcher to formulate specific expectations regarding the number and the nature of factors, and their correlations with each other. Furthermore, parameter estimates from one sample can be directly fit to data from other samples--the ultimate in elegant cross-validation. Such invariance tests are very important, since science is about finding relationships that replicate under stated conditions, and since statistical significance tests (even CFA tests) do not evaluate whether sample results occur in the population or are likely to occur in future samples (Thompson, 1996).

The present study was conducted to explore the use of item "testlets" or "parcels" when evaluating score validity using CFA. Some researchers define "testlets" as "a group of items related to a single content area that is developed as a unit" (Wainer & Kiely, 1978, p. 190), as for example when reading comprehension items are nested within the reading narratives with which the items are

associated. This was not the view of item "testlets" employed in the present study.

In the present study item "testlets" or item "parcels" were empirically created by combining items to create score aggregates that could then be subjected to CFA analysis. Such "testlets" can be useful in CFA, for two reasons.

First, the creation of such "testlets" enables the researcher to better meet the normal-distribution assumption of maximum-likelihood estimation of CFA factor pattern coefficients. CFA requires some rather strong distributional assumptions. It has long been recognized that item data can be combined so as to optimize the normality of data (e.g., Cattell, 1956; Cattell & Burdsal, 1975; Gorsuch, 1983, pp. 294-295; Gorsuch & Yagel, 1981). For example, item "testlets" can be created by pairing item responses with opposite skewness (e.g., most negatively skewed with most positively skewed within a given scale, etc.).

Second, combining items into "testlets" also results in more parsimonious model tests. One feature of this parsimony is that the rank of the estimated matrix of associations can be radically reduced. For example, if 78 items were the basis of analyses, initially 78 variances and 3003 (78 x 77 / 2 = 6006 / 2) unique covariances are estimated, and then the factor parameters to reproduce these coefficients are estimated. If the same 78 item responses are aggregated only into scores on 36 "doublets," initially only 36 variances and 630 (36 x 35 / 2 = 1260 / 2) unique covariances are estimated, and then the factor parameters to

reproduce these coefficients are estimated.

The number of factor parameters is also itself reduced. For example, presume for the model involving 78 items that each item is presumed to be "univocal" (i.e., to "speak" through one factor only) and that four uncorrelated factors are presumed. Here generally 156 parameters (i.e., 78 pattern/structure coefficients and 78 measurement error variances) would be estimated to reproduce the estimated 3081 (78 + 3003) population variances and covariances. For the model involving 36 "doublet testlets," generally 72 parameters (i.e., 36 pattern/structure coefficients and 36 measurement error variances) would be estimated to reproduce the estimated 666 (36 + 630) population variances and covariances.

Fitting more parsimonious models to reproduce fewer estimated population values in the matrix of associations leaves less room for sampling error to impact the estimation process. This in turn theoretically leads to results that better generalize.

Results

The heuristic example of "testlet" creation involves responses by 422 undergraduate students to the 78-item (previous) version of an instrument developed by the senior author, the Personal Preferences Self-Description Questionnaire (PPSDQ). The iterative development of the PPSDQ to date is detailed elsewhere (cf. Melancon & Thompson, 1994, 1996; Thompson & Melancon, 1995, 1996; Thompson & Stone, 1994). Here it was reasonable to analyze the correlation matrix, because in the measurement models no variable was allowed to be associated with more than one factor (Cudeck,

1989; Thompson & Daniel, 1996).

Appendix A presents descriptive statistics for each of the 78 items on this version of the PPSDQ. The data for many of these items have non-normal distributions (i.e., highly skewed, leptokurtic or platykurtic), as often occurs with attitudinal or behavioral item-level data.

Items within given scales were then aggregated into doublets with the view of maximizing normality of the "doublets" thus created. For example, item "V11676" was created for each of the 422 subjects by summing each subject's responses on items 16 and 76. Descriptive statistics for this iteration of item "testlets"/"parcels" are presented in Appendix B.

This process was continued through a total of five iterations, until the "testlets" described in Appendix E were created. Each of these "testlets" consisted of summated scores from responses to between six and 15 items on a given scale. As can be seen by comparing the descriptive statistics in the five appendices, each generation of "testlets" created data that were progressively more normally distributed.

The related maximum-likelihood model tests are reported in Tables 1 through 5. Each model posited the existence of the same four correlated factors. Thus, across analyses the so-called "measurement model" was modified since the measured variables were redefined each time, but the so-called "structural model" was the same in every analysis.

Discussion

The present study investigated the benefits of creating item "testlets" or "parcels" in the context of structural equation modeling confirmatory factor analysis. Item "testlets" were created to better meet the distributional assumptions of CFA. The benefit of using the packets is clear from the results.

Even though the measurement model was unaltered across the various iterations of analysis, the same structural model fit the data progressively more adequately as the distributions of data were progressively improved. For example, as reported in Tables 1 through 5, the adjusted goodness of fit statistics progressed from .682, .806, .860, .893, to .941. Similarly, the root mean square residuals progressed from .072, .068, .061, .052, to .031.

The creation of item "testlets" or "parcels" is recommended whenever data depart substantially from the distributional assumptions of confirmatory factor analysis. Augmenting analyses by creating "testlets" can provide more complete understanding regarding the quality of a structural model, by altering the measurement model while leaving the structural model intact.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: Author.

Cattell, R.B. (1956). Validation and intensification of the sixteen personality factor questionnaire. Journal of Clinical Psychology, 12, 205-214.

Cattell, R.B., & Burdsal, C.A. (1975). The radical parcel double factoring design: A solution to the item-vs-parcel controversy. Multivariate Behavioral Research, 10, 165-179.

Cronbach, L.J. (1989). Construct validation after thirty years. In R.L. Linn (Ed.), Intelligence: Measurement theory and public policy (pp. 147-171). Urbana: University of Illinois Press.

Cudeck, R. (1989). The analysis of correlation matrices using covariance structure models. Psychological Bulletin, 105, 317-327.

Gorsuch, R.L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Gorsuch, R.L., & Yagel, C. (1981, April). Item factor analysis. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Guilford, J.P. (1946). New standards for test evaluation. Educational and Psychological Measurement, 6, 427-439.

Joint Committee on Standards for Educational Evaluation. (1988). The personnel evaluation standards: How to assess systems for

evaluating educators. Thousand Oaks, CA: Sage.

Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards: How to assess evaluations of educational programs (2nd ed.). Thousand Oaks, CA: Sage.

Jöreskog, K.G., & Sörbom, D. (1989). LISREL 7: A guide to the program and applications (2nd ed.). Chicago: SPSS.

Melancon, J.G., & Thompson, B. (1994, November). An adjectival self-description checklist evaluating Myers-Briggs Type Indicator (MBTI) scores: Concurrent and construct score validity. Paper presented at the annual meeting of the Mid-South Educational Research Association, Nashville, TN. (ERIC Document Reproduction Service No. ED 379 339)

Melancon, J.G., & Thompson, B. (1996, April). Measurement of self-perceptions of Jungian psychological types. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. Review of Educational Research, 62, 229-258.

Moss, P.A. (1995). Themes and variations in validity theory. Educational Measurement: Issues and Practice, 14(2), 5-12.

Nunnally, J.C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Popper, K.R. (1962). Conjectures and refutations: The growth of scientific knowledge. New York: Harper & Row.

Shepard, L. A. (1993). Evaluating test validity. In L.

Darling-Hammond (Ed.), Review of research in education (Vol. 19, pp. 405-450). Washington, DC: American Educational Research Association.

Thompson, B. (1992). A partial test distribution for cosines among factors across samples. In B. Thompson (Ed.), Advances in social science methodology (Vol. 2, pp. 81-97). Greenwich, CT: JAI Press.

Thompson, B. (1994). Guidelines for authors. Educational and Psychological Measurement, 54(4), 837-847.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. Educational Researcher, 25(2), 26-30.

Thompson, B., & Daniel, L.G. (1996). Factor analytic evidence for the construct validity of scores: An historical overview and some guidelines. Educational and Psychological Measurement, 56, 213-224.

Thompson, B., & Melancon, J. (1995, January). Measurement integrity of scores from a self-description checklist evaluating Myers-Briggs Type Indicator (MBTI) types: A confirmatory factor analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas. (ERIC Document Reproduction Service No. ED 380 487)

Thompson, B., & Melancon, J.G. (1996, January). Measuring Jungian psychological types: Some confirmatory factor analyses. Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans.

Thompson, B., & Stone, E. (1994, January). Concurrent validity of scores from an adjectival self-description checklist in relation to Myers-Briggs Type Indicator (MBTI) scores. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX. (ERIC Document Reproduction Service No. ED 367 706)

Wainer, H., & Kiely, C. (1987). Item clusters and computerized adaptive testing: A case for testlets. Journal of Educational Measurement, 24, 185-202.

Table 1
Maximum Likelihood Estimates of CFA Factor Pattern Coefficients Using Item "Singlets" (n=422; v=78)

LAMBDA X

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
MIXERLON(1)	0.778	0.000	0.000	0.000
SOCPRIVA(1)	0.708	0.000	0.000	0.000
XINTREXT(1)	0.696	0.000	0.000	0.000
XSILENGA(1)	0.660	0.000	0.000	0.000
PERSNSHY(1)	0.760	0.000	0.000	0.000
XQUIETEX(1)	0.675	0.000	0.000	0.000
GREGARTI(1)	0.599	0.000	0.000	0.000
CONGRECL(1)	0.560	0.000	0.000	0.000
FRIEDIST(1)	0.678	0.000	0.000	0.000
XSOLIAMI(1)	0.594	0.000	0.000	0.000
EXUBSERE(1)	0.331	0.000	0.000	0.000
XSTILLAN(1)	0.461	0.000	0.000	0.000
XREFLECA(1)	0.268	0.000	0.000	0.000
APPROACH(1)	0.381	0.000	0.000	0.000
XTERSEWO(1)	0.276	0.000	0.000	0.000
SHY76(1)	0.586	0.000	0.000	0.000
XEASTA82(1)	0.608	0.000	0.000	0.000
TRADCREA(1)	0.000	0.602	0.000	0.000
PRECIMAG(1)	0.000	0.622	0.000	0.000
XINVENOR(1)	0.000	0.581	0.000	0.000
PLANVISI(1)	0.000	0.613	0.000	0.000
CONCLEXP(1)	0.000	0.538	0.000	0.000
XINSIGHT(1)	0.000	0.556	0.000	0.000
XDIVERCO(1)	0.000	0.501	0.000	0.000
REALINTU(1)	0.000	0.243	0.000	0.000
XDIVERPR(1)	0.000	0.589	0.000	0.000
XCONCEPR(1)	0.000	0.158	0.000	0.000
DIRECTIN(1)	0.000	0.350	0.000	0.000
PRACTHEO(1)	0.000	0.127	0.000	0.000
XVARIREP(1)	0.000	0.497	0.000	0.000
XINQUICR(1)	0.000	0.429	0.000	0.000
XLEFAC85(1)	0.000	0.395	0.000	0.000
INVENT88(1)	0.000	0.440	0.000	0.000
XMECHA91(1)	0.000	0.544	0.000	0.000
PERSPE94(1)	0.000	0.501	0.000	0.000
FACTCOMP(1)	0.000	0.000	0.669	0.000
XTENDERR(1)	0.000	0.000	0.624	0.000
XFEELTHI(1)	0.000	0.000	0.582	0.000
XKINDANA(1)	0.000	0.000	0.603	0.000
STRICTFO(1)	0.000	0.000	0.575	0.000
DISPASEM(1)	0.000	0.000	0.548	0.000
SKEPTRUS(1)	0.000	0.000	0.544	0.000
XEMPATHL(1)	0.000	0.000	0.454	0.000
LOGICHUM(1)	0.000	0.000	0.586	0.000
XLIGHTHE(1)	0.000	0.000	0.506	0.000

XGULLSUS(1)	0.000	0.000	0.352	0.000
XCARICOO(1)	0.000	0.000	0.481	0.000
XACCEPDI(1)	0.000	0.000	0.519	0.000
XRECEPTS(1)	0.000	0.000	0.420	0.000
EVALNONJ(1)	0.000	0.000	0.432	0.000
XSYMPATH(1)	0.000	0.000	0.265	0.000
JUSTHARM(1)	0.000	0.000	0.234	0.000
EVALOPEN(1)	0.000	0.000	0.481	0.000
PRINCIPL(1)	0.000	0.000	0.426	0.000
IMPERPER(1)	0.000	0.000	0.496	0.000
XSENSUAL(1)	0.000	0.000	0.285	0.000
XFLEXORG(1)	0.000	0.000	0.000	0.585
PROMPTFR(1)	0.000	0.000	0.000	0.607
XRANDSEQ(1)	0.000	0.000	0.000	0.586
TIMELYRE(1)	0.000	0.000	0.000	0.523
XIMPETTA(1)	0.000	0.000	0.000	0.470
XIMPULDE(1)	0.000	0.000	0.000	0.485
RESPADAP(1)	0.000	0.000	0.000	0.375
XCAREFRE(1)	0.000	0.000	0.000	0.559
XPLAN74(1)	0.000	0.000	0.000	0.548
HOLIDA75(1)	0.000	0.000	0.000	0.494
NOORGI77(1)	0.000	0.000	0.000	0.489
XSTFRE78(1)	0.000	0.000	0.000	0.540
XMALIS80(1)	0.000	0.000	0.000	0.383
PRESSU81(1)	0.000	0.000	0.000	0.340
GOFLOW83(1)	0.000	0.000	0.000	0.565
XHATER84(1)	0.000	0.000	0.000	0.294
XROUTI86(1)	0.000	0.000	0.000	0.527
CHANGE87(1)	0.000	0.000	0.000	0.448
LASTMI89(1)	0.000	0.000	0.000	0.408
XHAIMP90(1)	0.000	0.000	0.000	0.415
XONTIM92(1)	0.000	0.000	0.000	0.420
NOORDR93(1)	0.000	0.000	0.000	0.471

PHI

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
EXTRINTR	1.000			
SENSINTU	-0.412	1.000		
THINFEEL	-0.363	0.456	1.000	
JUDGPERC	-0.269	0.758	0.423	1.000

χ^2 WITH 2919 DEGREES OF FREEDOM = 6463.93 (P = .000)
GOODNESS OF FIT INDEX =0.699
ADJUSTED GOODNESS OF FIT INDEX =0.682
ROOT MEAN SQUARE RESIDUAL = 0.072

BEST COPY AVAILABLE

Table 2
Maximum Likelihood Estimates of CFA Factor Pattern Coefficients Using Item "Doublets" and "Singlets" (n=422; v=78)

LAMBDA X

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
V11676(2)	0.767	0.000	0.000	0.000
V12662(2)	0.721	0.000	0.000	0.000
V10166(2)	0.748	0.000	0.000	0.000
V15031(2)	0.484	0.000	0.000	0.000
V14670(2)	0.850	0.000	0.000	0.000
V13606(2)	0.667	0.000	0.000	0.000
V11182(2)	0.593	0.000	0.000	0.000
V15841(2)	0.772	0.000	0.000	0.000
V154(1)	0.557	0.000	0.000	0.000
V27194(2)	0.000	0.469	0.000	0.000
V26718(2)	0.000	0.484	0.000	0.000
V28588(2)	0.000	0.529	0.000	0.000
V20227(2)	0.000	0.523	0.000	0.000
V25112(2)	0.000	0.657	0.000	0.000
V26342(2)	0.000	0.696	0.000	0.000
V24759(2)	0.000	0.706	0.000	0.000
V29149(2)	0.000	0.653	0.000	0.000
V25507(2)	0.000	0.641	0.000	0.000
V30429(2)	0.000	0.000	0.622	0.000
V34372(2)	0.000	0.000	0.606	0.000
V33424(2)	0.000	0.000	0.485	0.000
V34509(2)	0.000	0.000	0.634	0.000
V35265(2)	0.000	0.000	0.719	0.000
V31948(2)	0.000	0.000	0.583	0.000
V36044(2)	0.000	0.000	0.729	0.000
V33039(2)	0.000	0.000	0.531	0.000
V35673(2)	0.000	0.000	0.707	0.000
V33225(2)	0.000	0.000	0.508	0.000
V364(1)	0.000	0.000	0.554	0.000
V49387(2)	0.000	0.000	0.000	0.615
V49278(2)	0.000	0.000	0.000	0.619
V47461(2)	0.000	0.000	0.000	0.669
V45753(2)	0.000	0.000	0.000	0.663
V48640(2)	0.000	0.000	0.000	0.695
V48077(2)	0.000	0.000	0.000	0.561
V47590(2)	0.000	0.000	0.000	0.596
V40589(2)	0.000	0.000	0.000	0.561
V48410(2)	0.000	0.000	0.000	0.629
V41720(2)	0.000	0.000	0.000	0.608
V48381(2)	0.000	0.000	0.000	0.567

PHI

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
EXTRINTR	1.000			
SENSINTU	-0.389	1.000		
THINFEEL	-0.334	0.434	1.000	
JUDGPERC	-0.244	0.751	0.403	1.000

χ^2 WITH 734 DEGREES OF FREEDOM = 1675.56 (P = .000)
GOODNESS OF FIT INDEX =0.826
ADJUSTED GOODNESS OF FIT INDEX =0.806
ROOT MEAN SQUARE RESIDUAL = 0.068

Table 3
Maximum Likelihood Estimates of CFA Factor Pattern Coefficients Using Item "Doublets," "Triplets" and "Quadruplets" ($\underline{n}$=422; $\underline{v}$=78)

LAMBDA X

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
V1001(4)	0.847	0.000	0.000	0.000
V1002(4)	0.822	0.000	0.000	0.000
V1003(4)	0.875	0.000	0.000	0.000
V1004(3)	0.627	0.000	0.000	0.000
V1005(2)	0.660	0.000	0.000	0.000
V2001(4)	0.000	0.783	0.000	0.000
V2002(4)	0.000	0.750	0.000	0.000
V2003(4)	0.000	0.678	0.000	0.000
V2004(4)	0.000	0.747	0.000	0.000
V2005(2)	0.000	0.664	0.000	0.000
V3001(4)	0.000	0.000	0.802	0.000
V3002(4)	0.000	0.000	0.620	0.000
V3003(4)	0.000	0.000	0.736	0.000
V3004(4)	0.000	0.000	0.719	0.000
V3005(3)	0.000	0.000	0.742	0.000
V3006(2)	0.000	0.000	0.719	0.000
V4001(4)	0.000	0.000	0.000	0.793
V4002(4)	0.000	0.000	0.000	0.752
V4003(4)	0.000	0.000	0.000	0.673
V4004(4)	0.000	0.000	0.000	0.807
V4005(4)	0.000	0.000	0.000	0.668
V4006(2)	0.000	0.000	0.000	0.672

PHI

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
EXTRINTR	1.000			
SENSINTU	-0.369	1.000		
THINFEEL	-0.313	0.417	1.000	
JUDGPERC	-0.239	0.728	0.410	1.000

χ^2 WITH 203 DEGREES OF FREEDOM = 549.16 (P = .000)
GOODNESS OF FIT INDEX =0.887
ADJUSTED GOODNESS OF FIT INDEX =0.860
ROOT MEAN SQUARE RESIDUAL = 0.061

Table 4
Maximum Likelihood Estimates of CFA Factor Pattern Coefficients Using "Testlets" Involving Four to Eight Items (n=422; v=78)

LAMBDA X

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
Z10106(6)	0.868	0.000	0.000	0.000
Z10207(8)	0.841	0.000	0.000	0.000
Z10304(4)	0.879	0.000	0.000	0.000
Z20108(8)	0.000	0.832	0.000	0.000
Z20206(6)	0.000	0.829	0.000	0.000
Z20304(4)	0.000	0.779	0.000	0.000
Z30108(8)	0.000	0.000	0.847	0.000
Z30207(7)	0.000	0.000	0.863	0.000
Z30306(6)	0.000	0.000	0.783	0.000
Z40108(8)	0.000	0.000	0.000	0.868
Z40208(8)	0.000	0.000	0.000	0.836
Z40306(6)	0.000	0.000	0.000	0.808

PHI

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
EXTRINTR	1.000			
SENSINTU	-0.363	1.000		
THINFEEL	-0.332	0.426	1.000	
JUDGPERC	-0.227	0.708	0.390	1.000

χ^2 WITH 48 DEGREES OF FREEDOM = 171.38 (P = .000)
GOODNESS OF FIT INDEX =0.934
ADJUSTED GOODNESS OF FIT INDEX =0.893
ROOT MEAN SQUARE RESIDUAL = 0.052

Table 5
Maximum Likelihood Estimates of CFA Factor Pattern Coefficients Using "Testlets" Involving Six to Fifteen Items ($\underline{n}$=422; $\underline{v}$=78)

LAMBDA X

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
Y10110(10)	0.822	0.000	0.000	0.000
Y10207(8)	0.946	0.000	0.000	0.000
Y20112(12)	0.000	0.892	0.000	0.000
Y20206(6)	0.000	0.824	0.000	0.000
Y30115(15)	0.000	0.000	0.935	0.000
Y30206(6)	0.000	0.000	0.770	0.000
Y40114(14)	0.000	0.000	0.000	0.970
Y40208(8)	0.000	0.000	0.000	0.821

PHI

	EXTRINTR	SENSINTU	THINFEEL	JUDGPERC
EXTRINTR	1.000			
SENSINTU	-0.392	1.000		
THINFEEL	-0.310	0.420	1.000	
JUDGPERC	-0.249	0.700	0.361	1.000

χ^2 WITH 14 DEGREES OF FREEDOM = 38.50 (P = .000)
GOODNESS OF FIT INDEX =0.977
ADJUSTED GOODNESS OF FIT INDEX =0.941
ROOT MEAN SQUARE RESIDUAL = 0.031

APPENDIX A

Descriptive Statistics for First Iteration of "Testlet" Creation

VARIABLE(v)	MEAN	ST. DEV.	SKEWNESS	KURTOSIS	MINIMUM	FREQ.	MAXIMUM	FREQ.
MIXERLON(1)	3.064	1.662	0.531	-0.586	1.000	89	7.000	14
SOCPRIVA(1)	3.592	1.754	0.197	-1.001	1.000	55	7.000	22
XINTREXT(1)	3.355	1.695	0.338	-0.745	1.000	67	7.000	20
XSILENGA(1)	3.848	1.555	0.026	-0.674	1.000	31	7.000	19
PERSNSHY(1)	2.953	1.754	0.699	-0.532	1.000	105	7.000	18
XQUIETEX(1)	3.175	1.614	0.622	-0.495	1.000	54	7.000	15
GREGARTI(1)	3.476	1.395	0.178	-0.481	1.000	32	7.000	6
CONGRECL(1)	3.050	1.293	0.344	-0.056	1.000	50	7.000	5
FRIEDIST(1)	2.230	1.425	1.352	1.547	1.000	168	7.000	9
XSOLIAMI(1)	3.192	1.373	0.387	-0.212	1.000	43	7.000	7
EXUBSERE(1)	3.941	1.527	-0.032	-0.675	1.000	26	7.000	16
XSTILLAN(1)	2.922	1.387	0.575	-0.358	1.000	58	7.000	3
XREFLECA(1)	3.467	1.583	0.391	-0.638	1.000	37	7.000	16
APPROACH(1)	3.344	1.751	0.511	-0.829	1.000	54	7.000	22
XTERSEWO(1)	3.763	1.428	0.108	-0.364	1.000	21	7.000	15
SHY76(1)	4.258	2.025	-0.247	-1.226	1.000	58	7.000	67
XEASTA82(1)	3.488	2.006	0.262	-1.285	1.000	92	7.000	32
TRADCREA(1)	4.855	1.768	-0.520	-0.760	1.000	19	7.000	91
PRECIMAG(1)	4.765	1.632	-0.524	-0.587	1.000	15	7.000	59
XINVENOR(1)	4.093	1.724	-0.166	-0.972	1.000	38	7.000	27
PLANVISI(1)	4.694	1.535	-0.495	-0.548	1.000	11	7.000	39
CONCLEXP(1)	5.227	1.295	-0.792	0.326	1.000	3	7.000	59
XINSIGHT(1)	4.630	1.579	-0.335	-0.684	1.000	11	7.000	52
XDIVERCO(1)	4.652	1.648	-0.292	-0.973	1.000	8	7.000	58
REALINTU(1)	3.675	1.590	0.073	-0.956	1.000	35	7.000	11
XDIVERPR(1)	4.493	1.597	-0.163	-0.849	1.000	10	7.000	51
XCONCEPR(1)	3.310	1.506	0.485	-0.459	1.000	39	7.000	11
DIRECTIN(1)	4.530	1.463	-0.159	-0.494	1.000	10	7.000	42
PRACTHEO(1)	3.311	1.416	0.436	-0.253	1.000	38	7.000	7
XVARIREP(1)	5.329	1.492	-0.727	-0.203	1.000	4	7.000	113
XINQUICR(1)	4.721	1.428	-0.341	-0.562	1.000	5	7.000	41
XLEFAC85(1)	3.398	1.805	0.367	-0.931	1.000	72	7.000	23
INVENT88(1)	5.199	1.420	-0.730	0.290	1.000	8	7.000	85
XMECHA91(1)	4.663	1.720	-0.267	-0.927	1.000	15	7.000	75
PERSPE94(1)	5.786	1.251	-0.994	0.650	1.000	1	7.000	157
FACTCOMP(1)	4.789	1.567	-0.618	-0.386	1.000	14	7.000	47
XTENDERR(1)	4.368	1.612	-0.198	-0.853	1.000	15	7.000	37
XFEELTHI(1)	4.268	1.750	-0.096	-1.021	1.000	22	7.000	49
XKINDANA(1)	4.858	1.651	-0.605	-0.486	1.000	16	7.000	68
STRICTFO(1)	5.140	1.468	-0.846	0.086	1.000	7	7.000	66
DISPASEM(1)	5.511	1.338	-0.819	0.068	1.000	3	7.000	111
SKEPTRUS(1)	4.555	1.651	-0.338	-0.871	1.000	14	7.000	46
XEMPATHL(1)	3.393	1.537	0.448	-0.680	1.000	33	7.000	10
LOGICHUM(1)	4.403	1.701	-0.264	-0.922	1.000	21	7.000	43
XLIGHTHE(1)	4.908	1.494	-0.507	-0.395	1.000	7	7.000	61
XGULLSUS(1)	3.413	1.523	0.387	-0.441	1.000	40	7.000	15
XCARICOO(1)	5.374	1.427	-0.832	0.167	1.000	4	7.000	104
XACCEPDI(1)	5.152	1.540	-0.738	-0.117	1.000	9	7.000	90
XRECEPTS(1)	4.261	1.685	-0.215	-1.007	1.000	21	7.000	30
EVALNONJ(1)	4.033	1.632	0.052	-0.881	1.000	23	7.000	29
XSYMPATH(1)	3.583	1.683	0.296	-0.829	1.000	43	7.000	22
JUSTHARM(1)	4.002	1.878	-0.174	-1.115	1.000	60	7.000	34
EVALOPEN(1)	4.664	1.622	-0.475	-0.668	1.000	16	7.000	44
PRINCIPL(1)	4.755	1.734	-0.539	-0.725	1.000	19	7.000	67
IMPERPER(1)	5.557	1.329	-1.238	1.755	1.000	9	7.000	106
XSENSUAL(1)	4.602	1.662	-0.318	-0.810	1.000	15	7.000	57

XFLEXORG(1)	4.071	1.723	-0.094	-0.948	1.000	34	7.000	32
PROMPTFR(1)	4.385	1.871	-0.255	-1.032	1.000	37	7.000	65
XRANDSEQ(1)	3.938	1.621	0.093	-0.906	1.000	21	7.000	24
TIMELYRE(1)	4.370	1.715	-0.288	-0.934	1.000	25	7.000	39
XIMPETTA(1)	3.537	1.430	0.342	-0.294	1.000	26	7.000	13
XIMPULDE(1)	4.358	1.533	-0.083	-0.773	1.000	9	7.000	37
RESPADAP(1)	3.621	1.768	0.121	-1.104	1.000	57	7.000	18
XCAREFRE(1)	4.392	1.700	-0.216	-0.849	1.000	24	7.000	49
XPLAN74(1)	2.945	1.642	0.663	-0.470	1.000	93	7.000	9
HOLIDA75(1)	3.867	1.710	0.139	-0.987	1.000	30	7.000	30
NOORGI77(1)	4.375	1.887	-0.189	-1.147	1.000	31	7.000	68
XSTFRE78(1)	4.813	1.558	-0.435	-0.685	1.000	7	7.000	58
XMALIS80(1)	3.908	1.923	0.152	-1.138	1.000	49	7.000	56
PRESSU81(1)	3.886	1.994	-0.072	-1.307	1.000	76	7.000	38
GOFLOW83(1)	4.145	1.719	-0.053	-0.949	1.000	29	7.000	39
XHATER84(1)	3.611	1.835	0.098	-1.127	1.000	72	7.000	23
XROUTI86(1)	3.727	1.590	0.183	-0.682	1.000	35	7.000	19
CHANGE87(1)	5.396	1.420	-0.808	0.161	1.000	5	7.000	111
LASTMI89(1)	4.071	1.839	-0.119	-1.096	1.000	45	7.000	41
XHAIMP90(1)	4.138	1.578	-0.124	-0.719	1.000	23	7.000	26
XONTIM92(1)	2.599	1.613	0.769	-0.423	1.000	147	7.000	6
NOORDR93(1)	2.559	1.526	0.915	0.126	1.000	129	7.000	7

APPENDIX B

Descriptive Statistics for Second Iteration of "Testlet" Creation

VARIABLE(v)	MEAN	ST. DEV.	SKEWNESS	KURTOSIS	MINIMUM	FREQ.	MAXIMUM	FREQ.
V11676(2)	6.488	2.818	0.209	-0.458	2.000	39	14.000	5
V12662(2)	6.893	2.590	0.207	-0.451	2.000	16	13.000	10
V10166(2)	7.024	2.783	0.386	-0.552	2.000	11	14.000	6
V15031(2)	6.685	2.129	0.227	-0.062	2.000	8	13.000	2
V14670(2)	6.540	2.568	0.410	-0.251	2.000	16	14.000	3
V13606(2)	6.936	2.815	0.353	-0.496	2.000	17	14.000	7
V11182(2)	6.955	2.753	0.199	-0.703	2.000	17	14.000	1
V15841(2)	6.547	2.581	0.359	-0.174	2.000	20	14.000	5
V154(1)	3.050	1.293	0.344	-0.056	1.000	50	7.000	5
V27194(2)	9.097	1.964	0.035	0.253	2.000	1	14.000	5
V26718(2)	8.539	1.903	0.104	0.040	3.000	1	14.000	2
V28588(2)	8.597	2.569	-0.111	-0.302	2.000	4	14.000	7
V20227(2)	9.005	2.263	-0.217	-0.239	2.000	1	14.000	5
V25112(2)	9.295	2.382	-0.154	-0.344	2.000	1	14.000	16
V26342(2)	9.348	2.807	-0.308	-0.368	2.000	5	14.000	30
V24759(2)	8.787	2.724	-0.254	-0.492	2.000	6	14.000	13
V29149(2)	9.384	2.400	-0.115	-0.561	2.000	1	14.000	15
V25507(2)	9.282	2.576	-0.202	-0.643	3.000	5	14.000	13
V30429(2)	8.950	2.192	-0.181	0.229	2.000	3	14.000	3
V34372(2)	8.553	2.282	-0.107	0.139	2.000	2	14.000	7
V33424(2)	8.957	2.319	-0.111	0.088	2.000	2	14.000	14
V34509(2)	9.544	2.275	-0.293	0.004	2.000	2	14.000	14
V35265(2)	9.419	2.583	-0.338	-0.138	2.000	5	14.000	24
V31948(2)	8.791	2.574	-0.382	-0.125	2.000	8	14.000	5
V36044(2)	9.226	2.735	-0.426	-0.235	2.000	5	14.000	19
V33039(2)	9.016	2.696	-0.403	-0.230	2.000	8	14.000	11
V35673(2)	9.310	2.517	-0.199	-0.420	2.000	2	14.000	18
V33225(2)	9.265	2.436	-0.347	0.053	2.000	3	14.000	15
V364(1)	4.555	1.651	-0.338	-0.871	1.000	14	7.000	46
V49387(2)	7.955	2.222	0.130	0.217	2.000	3	14.000	5
V49278(2)	7.411	2.474	0.171	-0.376	2.000	4	14.000	2
V47461(2)	7.315	2.704	0.016	-0.572	2.000	14	14.000	2
V45753(2)	7.922	2.687	0.010	-0.424	2.000	7	14.000	10
V48640(2)	8.119	2.579	-0.066	-0.601	2.000	4	14.000	3
V48077(2)	8.283	2.924	-0.052	-0.711	2.000	7	14.000	10
V47590(2)	8.005	2.541	-0.055	-0.483	2.000	9	14.000	1
V40589(2)	7.692	2.675	0.034	-0.484	2.000	10	14.000	6
V48410(2)	7.683	2.591	0.013	-0.409	2.000	9	14.000	6
V41720(2)	8.296	2.666	-0.045	-0.511	2.000	3	14.000	9
V48381(2)	8.031	2.942	-0.007	-0.709	2.000	12	14.000	10

APPENDIX C
Descriptive Statistics for Third Iteration of "Testlet" Creation

VARIABLE(v)	MEAN	ST. DEV.	SKEWNESS	KURTOSIS	MINIMUM	FREQ.	MAXIMUM	FREQ.
V1001(4)	13.495	4.567	0.244	-0.524	4.000	3	26.000	1
V1002(4)	13.917	4.780	0.351	-0.298	4.000	2	27.000	2
V1003(4)	13.035	4.776	0.280	-0.252	4.000	9	28.000	1
V1004(3)	9.734	2.777	0.196	0.149	3.000	6	19.000	1
V1005(2)	6.936	2.815	0.353	-0.496	2.000	17	14.000	7
V2001(4)	17.887	3.792	-0.158	-0.257	6.000	1	28.000	1
V2002(4)	17.884	3.817	-0.106	-0.374	8.000	2	28.000	2
V2003(4)	17.602	3.835	-0.203	-0.241	7.000	3	26.000	3
V2004(4)	18.666	4.222	-0.061	-0.485	6.000	1	28.000	4
V2005(2)	9.295	2.382	-0.154	-0.344	2.000	1	14.000	16
V3001(4)	17.779	4.240	-0.408	-0.063	5.000	1	28.000	1
V3002(4)	17.973	4.096	-0.318	0.081	5.000	1	28.000	2
V3003(4)	17.742	3.905	-0.301	0.232	4.000	1	28.000	1
V3004(4)	18.576	4.116	-0.150	-0.030	5.000	1	28.000	4
V3005(3)	14.098	3.228	-0.356	0.134	3.000	2	21.000	8
V3006(2)	9.419	2.583	-0.338	-0.138	2.000	5	14.000	24
V4001(4)	15.531	4.244	0.002	-0.452	5.000	1	27.000	1
V4002(4)	15.960	3.864	0.058	-0.139	5.000	2	27.000	1
V4003(4)	15.975	4.542	-0.026	-0.434	5.000	1	27.000	3
V4004(4)	15.611	4.379	-0.026	-0.488	5.000	4	26.000	1
V4005(4)	15.713	4.746	0.118	-0.289	4.000	3	28.000	3
V4006(2)	7.922	2.687	0.010	-0.424	2.000	7	14.000	10

APPENDIX D
Descriptive Statistics for Fourth Iteration of "Testlet" Creation

VARIABLE(v)	MEAN	ST. DEV.	SKEWNESS	KURTOSIS	MINIMUM	FREQ.	MAXIMUM	FREQ.
Z10106(6)	20.431	6.581	0.239	-0.562	7.000	3	38.000	2
Z10207(8)	23.651	6.690	0.303	-0.071	8.000	2	46.000	1
Z10304(4)	13.035	4.776	0.280	-0.252	4.000	9	28.000	1
Z20108(8)	36.268	6.956	-0.044	-0.284	15.000	1	53.000	1
Z20206(6)	27.179	5.396	-0.118	-0.408	12.000	1	42.000	1
Z20304(4)	17.887	3.792	-0.158	-0.257	6.000	1	28.000	1
Z30108(8)	36.355	7.407	-0.277	0.164	11.000	1	54.000	1
Z30207(7)	31.840	6.209	-0.394	0.332	9.000	1	49.000	1
Z30306(6)	27.393	5.677	-0.354	0.462	7.000	1	42.000	1
Z40108(8)	31.325	7.963	0.113	-0.335	11.000	1	52.000	3
Z40208(8)	31.935	7.374	0.154	-0.190	11.000	1	53.000	1
Z40306(6)	23.452	6.112	0.084	-0.461	8.000	1	40.000	1

APPENDIX E
Descriptive Statistics for Fifth Iteration of "Testlet" Creation

UNIVARIATE SUMMARY STATISTICS FOR CONTINUOUS VARIABLES

VARIABLE (v)	MEAN	ST. DEV.	SKEWNESS	KURTOSIS	MINIMUM	FREQ.	MAXIMUM	FREQ.
Y10110(10)	33.466	10.695	0.327	-0.523	11.000	1	63.000	1
Y10207(8)	23.651	6.690	0.303	-0.071	8.000	2	46.000	1
Y20112(12)	54.155	9.842	-0.053	-0.274	24.000	1	80.000	1
Y20206(6)	27.179	5.396	-0.118	-0.408	12.000	1	42.000	1
Y30115(15)	68.195	12.674	-0.359	0.323	21.000	1	101.000	1
Y30206(6)	27.393	5.677	-0.354	0.462	7.000	1	42.000	1
Y40114(14)	55.388	12.254	0.223	-0.453	24.000	1	91.000	1
Y40208(8)	31.325	7.963	0.113	-0.335	11.000	1	52.000	3

TM026107

U.S. DEPARTMENT OF EDUCATION

Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: USING ITEM "TESTLETS"/"PARCELS" IN CONFIRMATORY FACTOR ANALYSIS: AN EXAMPLE USING THE PPSDQ-78

Author(s): BRUCE THOMPSON and JANET G. MELANCON

Corporate Source:

Publication Date: 11/7/96

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here

Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

BRUCE THOMPSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here

Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature:	Position: PROFESSOR
Printed Name: BRUCE THOMPSON	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1335
	Date: 9/28/96

OVER

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC , *or*, if you wish ERIC to cite the availability of this document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents which cannot be made available through EDRS).

Publisher/Distributor:	
Address:	
Price Per Copy:	Quantity Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name and address of current copyright/reproduction rights holder:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

If you are making an unsolicited contribution to ERIC, you may return this form (and the document being contributed) to:

ERIC Facility
1301 Piccard Drive, Suite 300
Rockville, Maryland 20850-4305
Telephone: (301) 258-5500